American Journal of Epidemiology

Vol. 150, No. 12
Printed in U.S.A.

Effect of Smoking on the Body Mass Index-Mortality Relation: Empirical Evidence from 15 Studies

The BMI in Diverse Populations Collaborative Group*

The authors examined the impact of smoking status on the relation between body mass index (weight (kg)/height (m)^2) and mortality across a group of 15 diverse observational studies. The studies included a heterogeneous sample of national samples, cohort studies with mortality follow-up, and clinical trials. Consideration of the data according to natural strata resulted in the formation of 42 analytic cohorts. The authors examined survival through the end of follow-up for each study, as influenced by body mass index, age, and current smoking status at baseline, using a proportional hazards model to describe the relation between body mass index and mortality with control for age and smoking status. In this paper, the authors demonstrate that the estimated body mass index of minimum mortality changes when data are analyzed while ignoring smoking status; but they also demonstrate through a simulation study that eliminating smokers from the data sets prior to analysis produces results similar to those expected from the elimination of numerically similar random proportions of the data sets prior to analysis. Based on the results of these analyses, the authors find no support for the commonly held practice of eliminating smokers from a data set prior to examining the body mass index-mortality relation. *Am J Epidemiol* 1999;150:1297–308.

body mass index; cohort studies; epidemiologic methods; mortality; smoking

Many researchers have examined the consequences of obesity in order to ascertain the health effects of excess weight. One important component of this research has been the determination of the range of body composition associated with minimal morbidity and mortality. Longitudinal studies provide evidence concerning the parameters associated with minimal risk, but while findings often appear to reinforce one another, interpretations vary. Differences in study designs and methods of analysis may explain the divergent outcomes.

For more than half a century, the insurance industry has shown weight (adjusted for height) to be a predictor of mortality, and beginning in 1955 the Metropolitan Life Insurance Company established tables of desirable weights by height for general use. Epidemiologists now prefer to use a single measure called the body mass index (BMI), defined as weight (in kilograms) divided by height (in meters) squared, to determine obesity. The latest federal guidelines are based on this calculation. (The most recent recommendations state that individuals with a BMI greater than 25 are at risk of developing diseases associated with overweight and obesity (1).)

Epidemiologists have examined various health risk factors to determine whether they affect the relation between BMI and morbidity and mortality. One such risk factor discussed extensively is cigarette smoking; but while all investigators acknowledge the need to determine the impact of smoking on this relation, the appropriate methods for achieving this aim remain subject to debate. The question is whether the shape of the curve describing mortality in terms of BMI (i.e., the parameters of that curve) differs for smokers and nonsmokers. In technical terms, this problem is formulated by asking whether there is an interaction between cigarette smoking and weight or whether cigarette smoking is a confounder in this relation. Reports on this subject have been inconsistent.

We conducted an analysis of person-level data from 15 studies (2–19) to examine the effect of smoking on the relation between BMI and mortality by addressing the following questions: Is the shape of the curve for mortality by BMI different for smokers and nonsmokers? Does failure to consider cigarette smoking as a confounder alter the point of minimum mortality? Does restricting an analysis to nonsmokers produce results that differ from those obtained when one randomly reduces the size of the analytic sample?

Received for publication October 9, 1998, and accepted for publication March 24, 1999.

Abbreviations: BMI, body mass index; LBMI, lean body mass index; SE, standard error.

*A complete list of coinvestigators is provided in the Acknowledgment.

MATERIALS AND METHODS

Studies

To obtain a heterogeneous sample, we included in our analyses national samples with mortality follow-up, cohort studies with mortality follow-up, and clinical trials. The national samples included the pooled National Health Interview Survey data for 1987–1990 (2) and data from the First (3) and Second (4) National Health and Nutrition Examination Surveys. Cohort studies included the Framingham Heart Study (5), the Puerto Rico Heart Health Program (6), the Tecumseh Community Health Study (7), the Yugoslavia Cardiovascular Disease Study (8), the Honolulu Heart Program (9), the Scottish Collaborative Study (10), the Renfrew and Paisley Survey (11), the Israeli Ischemic Heart Disease Study (12), the Glostrup Cohort (13–16), and the Lipid Research Clinics Program Prevalence Study (17). Clinical trials included the Hypertension Detection and Follow-up Program (18) and the Multiple Risk Factor Intervention Trial (19). Baseline information began in 1948 with the Framingham Heart Study and continued through 1990 with the National Health Interview Survey. Data samples included people of both sexes and White and Black subgroups. Follow-up varied from 6 years for the National Health Interview Survey to 30 years for the Framingham cohort (table 1). Our final analytic sample was taken from 15 studies, producing 250,182 participants and 38,532 deaths (table 1).

The studies analyzed often contained subgroups based on factors such as sex, race/ethnicity, area of residence, and treatment status (in the case of clinical trials). These subgroups were analyzed separately if they met the criterion of having at least 15 deaths for each parameter included in our most complex model. The Lipid Research Clinics study and the Multiple Risk Factor Intervention Trial contained too few African Americans to permit separate analyses. In all, 42 cohorts were analyzed (table 2).

We included in our analyses participants with known values for age, smoking status, and BMI, which necessitated only a small number of exclusions. The studies exhibited a broad distribution span for BMI levels, ranging from an average BMI of 22.2 for the rural Yugoslavian male cohort to 30.2 for African-American women from the referred-care group of the Hypertension Detection and Follow-up Program. Although average age did not differ greatly among the cohorts, the age ranges showed substantial variation (tables 1 and 2). When we classified only current smokers as smokers, we found their proportion to vary from slightly more than 25 percent among white females in the National Health Interview Survey to over 71 percent among rural Yugoslavian men (table 2). Since we did not have information on number of cigarettes smoked per day for all cohorts, we simply separated participants into smokers and nonsmokers.

Detailed descriptions of the studies included in our analysis are provided in the Appendix.

Statistical methods

To model the relation between BMI and mortality, we used a single algorithmic approach to subject individual cohorts to a uniform method of analysis. A complete report on the development of our methods has been published elsewhere (20).

TABLE 1. Studies included in a person-level meta-analysis of body mass index and mortality

	Baseline year(s)	No. of observations	No. of deaths	No. of years of follow-up	Age range (years)
National Health Interview Survey	1987–1990	121,208	9,577	6	18–90
NHANES* I	1971–1975	12,730	4,016	19	24–77
NHANES II	1976–1980	9,064	2,106	14	30–75
Framingham Heart Study	1948–1951	5,163	1,964	30	28–62
Puerto Rico Heart Health Program	1965	9,776	1,726	15	35–79
Tecumseh Community Health Study	1959–1960	4,580	959	18	18–91
Yugoslavia Cardiovascular Disease Study	1964	6,450	1,337	16	34–62
Honolulu Heart Program	1965	8,005	2,466	21	45–68
Scottish Collaborative Study	1970–1973	7,008	2,079	22	21–75
Renfrew and Paisley Survey	1972–1976	15,394	4,443	17	45–64
Israeli Ischemic Heart Disease Study	1963	10,027	3,464	22	39–74
Glostrup Cohort	1977–1992	10,135	1,100	9	30–80
Lipid Research Clinics Program	1971–1976	8,175	947	12	30–97
Hypertension Detection and Follow-up Program	1973–1974	10,908	1,415	8	30–69
Multiple Risk Factor Intervention Trial	1973–1975	11,559	933	10	35–58
Total		250,182	38,532		

* NHANES, National Health and Nutrition Examination Survey.

With few exceptions, investigators report the relation between BMI and mortality to be nonmonotonic, in that excess mortality is associated with both high and low BMIs. To model such a relation and to account for asymmetry, we introduced a transformation of BMI into "lean" BMI (1/BMI) as the independent variable in a proportional hazards model. Previous analyses had suggested that the relation between BMI and mortality was asymmetric (20), so we sought a transformation to normality to account for this. In seeking a transformation to normality, we were influenced by Cornfield et al. (21), who demonstrated that if one assumes normality for both deaths and nondeaths and if, in addition, the variance differs for the two distributions, a quadratic model is necessary to describe the relation adequately in a logistic regression. This transformation was suggested by Nevill and Holder (22), who termed 1/BMI the "lean" BMI (LBMI). They demonstrated that LBMI was normally distributed using data from the Allied Dunbar National Fitness Study and that it was more closely related to percentage of body fat than BMI.

We used a proportional hazards model to describe the relation between LBMI and mortality, fitting each cohort to determine which model best described the relation. We assumed the following hazard:

$$\lambda(t) = \lambda_0(t) \times \exp\{\beta_1(\text{age}) + \beta_2(\text{smoking}) + \beta_3(\text{LBMI}) + \beta_4(\text{LBMI}^2) + \beta_5(\text{smoking} \times \text{LBMI}) + \beta_6(\text{smoking} \times \text{LBMI}^2)\},$$

where smoking is an indicator variable, coded 1 for smokers and 0 for nonsmokers, and the β's comprise a vector of unknown parameters to be estimated using maximum likelihood (given a cohort with information on the characteristics of participants and deaths during follow-up). For our purposes, we assumed β_1 and β_2 to be nonzero and developed an algorithm to determine which of the remaining parameters were necessary to best fit the data for each cohort.

We used a hierarchical modeling algorithm to decide which of four models was best supported by the data in each cohort. The four models considered were as follows.

Model 1. $\beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$.

This model assumes that only age and smoking are significant predictors of mortality.

Model 2. $\beta_3 \neq 0; \beta_4 = \beta_5 = \beta_6 = 0$.

This model assumes that LBMI is related to mortality in a monotonic fashion. The direction of the relation depends on the sign of the β_3.

Model 3. $\beta_3 \neq 0; \beta_4 \neq 0; \beta_5 = \beta_6 = 0$.

This model assumes a quadratic relation between LBMI and mortality.

Model 4. $\beta_3 \neq 0; \beta_4 \neq 0; \beta_5 \neq 0; \beta_6 \neq 0$.

The final model assumes that the basic relation between LBMI and mortality differs for smokers and nonsmokers.

We began by assuming that model 1 was the correct model and then examining whether each successively more complex model added significant information. The decision that a more complex model added significant information to a less complex model was made on the basis of a generalized likelihood ratio statistic. If model 1 was found to fit best, we concluded that the cohort showed no relation between BMI and mortality. If model 2 was selected, we determined that there was either a direct relation or an inverse relation, depending on the sign of β_3. If the third model fitted best, we considered the relation to be quadratic; and for model 4, we classified it as exhibiting interaction.

For those cases in which the quadratic model was deemed most appropriate, the estimated BMI of minimum mortality (BMI_{min}) and its standard error were calculated using standard statistical procedures (20):

$$\text{BMI}_{\text{min}} = -\left(\frac{2\hat{\beta}_4}{\hat{\beta}_3}\right),$$

and the variance of BMI_{min} was estimated using the delta method:

$$\text{Var}(\text{BMI}_{\text{min}}) = 4\left(\frac{\hat{\beta}_4^2}{\hat{\beta}_3^2}\right)\left[\frac{\text{Var}(\hat{\beta}_3, \hat{\beta}_4)}{\hat{\beta}_3^2} - 2\frac{\text{Cov}(\hat{\beta}_3, \hat{\beta}_4)}{\hat{\beta}_3\hat{\beta}_4} + \frac{\text{Var}(\hat{\beta}_4)}{\hat{\beta}_4}\right].$$

We reanalyzed the relation between BMI and mortality for each cohort after eliminating smokers from the data set. To determine whether this exclusion would produce significant changes in the relation, we performed an approximate randomization test (23). This test randomly assigned smoking status to individuals from each cohort. The number of random participants designated smokers equaled the number of actual smokers within each cohort. These designated smokers

TABLE 2. Analytic cohorts and results used in a meta-analysis relating body mass index to mortality

No. of observations	No. of deaths	Body mass index*		Age (years)		% smokers	Best model (relation)	BMI_{min}†	Standard error	Subjects
		Mean	SD†	Mean	SD					
				National Health Interview Survey						
59,036	4,228	24.3	5.0	46.7	18.9	25.4	Quadratic	23.8	0.35	White females
44,866	3,829	25.6	3.9	44.3	17.4	29.5	Interaction			White males
11,171	779	26.9	6.0	43.4	17.8	28.2	Quadratic	26.3	0.72	Black females
6,135	741	26.0	4.6	44.4	17.4	37.1	Quadratic	26.9	0.96	Black males
				First National Health and Nutrition Examination Survey						
6,466	1,538	25.3	5.4	47.7	15.4	30.8	Quadratic	23.4	0.45	White females
4,430	1,774	25.8	4.0	51.1	15.1	41.4	Quadratic	23.5	0.41	White males
1,184	373	27.9	6.7	47.7	15.2	36.4	Quadratic	26.7	1.08	Black females
650	331	25.7	5.0	53.0	14.9	47.5	Quadratic	25.5	0.61	Black males
				Second National Health and Nutrition Examination Survey						
4,267	775	25.9	5.6	54.5	13.4	28.4	Quadratic	25.2	0.83	White females
3,804	1,102	25.9	3.9	54.4	13.2	35.8	None			White males
547	97	28.7	6.4	53.6	13.2	28.2	Quadratic	26.3	1.56	Black females
446	132	25.8	4.7	54.2	13.7	47.1	None			Black males
				Framingham Heart Study						
2,850	886	25.3	4.7	44.1	8.5	41.1	Quadratic	23.5	0.59	White females
2,313	1,078	25.7	3.5	44.1	8.6	64.6	Quadratic	23.0	0.53	White males
				Puerto Rico Heart Health Program						
6,806	1,234	26.0	4.1	54.2	6.4	41.2	Quadratic	23.6	0.87	Urban males
2,970	492	23.3	3.5	55.0	7.0	49.6	Quadratic	22.9	0.60	Rural males
				Tecumseh Community Health Study						
2,389	388	25.2	5.2	41.0	15.7	35.1	Quadratic	22.1	1.21	White females
2,191	571	25.5	3.8	41.5	14.9	60.2	Quadratic	23.7	0.78	White males
				Yugoslavia Cardiovascular Disease Study						
3,548	683	24.3	3.6	45.2	7.5	68.5	Quadratic	24.3	0.61	Urban males
2,902	654	22.2	2.5	46.7	7.8	71.6	Quadratic	24.6	1.09	Rural males
				Honolulu Heart Program						
8,005	2,466	23.8	3.2	54.4	5.6	43.7	Quadratic	20.9	0.36	Males
				Scottish Collaborative Study						
6,006	1,904	25.1	3.1	47.6	7.3	55.1	Quadratic	22.5	0.52	Males
1,002	175	24.7	3.8	47.1	6.7	58.8	None			Females
				Renfrew and Paisley Survey						
7,055	2,549	25.9	3.4	54.1	5.6	56.6	Quadratic	24.1	0.71	Males
8,339	1,894	25.8	4.5	54.4	5.6	46.7	Interaction			Females

Table continues

were then excluded from the analysis. By repeating this process 1,000 times, we were able to judge whether the elimination of smokers produced results that differed significantly from those due to random exclusions.

RESULTS

Table 2 shows the results obtained from modeling the relation between BMI, age, smoking, and mortality. For 30 of the 42 analytic cohorts, we found the quadratic relation between BMI and mortality to be the most appropriate, while controlling for smoking. For eight of the remaining cohorts, we found no relation between BMI and mortality. One remaining cohort showed a direct relation between BMI and mortality, and the other three showed interactive relations in which the relation between BMI and mortality differed for smokers and nonsmokers.

TABLE 2. Continued

No. of observations	No. of deaths	Body mass index* Mean	Body mass index* SD†	Age (years) Mean	Age (years) SD	% smokers	Best model (relation)	BMI_{min}†	Standard error	Subjects
				Israeli Ischemic Heart Disease Study						
10,027	3,464	25.7	3.3	49.4	6.9	51.9	Interaction			Males
				Glostrup Cohort						
5,059	437	24.2	4.3	48.9	13.7	46.8	Quadratic	24.8	0.83	Females
5,076	663	25.4	3.5	47.6	7.3	56.8	Quadratic	24.8	0.86	Males
				Lipid Research Clinics Program, random sample						
2,198	191	24.5	4.5	48.2	12.4	29.8	None			Females
2,436	286	26.3	3.4	46.9	11.7	35.3	Quadratic	25.2	0.88	Males
				Lipid Research Clinics Program, hyperlipidemic sample						
1,532	185	25.8	5.0	49.7	13.0	36.8	None			Females
2,009	285	27.3	3.5	46.2	11.4	39.0	Quadratic	26.2	0.79	Males
				Hypertension Detection and Follow-up Program, stepped care						
1,178	103	28.2	6.3	52.7	9.2	29.8	Direct			White females
1,893	223	28.1	4.3	50.7	9.6	35.1	Quadratic	29.4	2.53	White males
1,339	133	29.5	7.0	49.6	10.2	37.8	None			Black females
1,064	196	27.1	5.2	50.2	10.3	55.5	Quadratic	29.1	2.08	Black males
				Hypertension Detection and Follow-up Program, referred care						
1,145	118	28.2	6.4	52.3	9.3	31.2	None			White females
1,857	235	28.2	4.3	50.8	9.6	35.2	Quadratic	26.7	1.34	White males
1,352	167	30.2	7.0	49.2	10.1	36.5	Quadratic	33.1	3.36	Black females
1,080	240	27.1	4.9	50.9	10.0	56.5	Quadratic	30.5	2.94	Black males
				Multiple Risk Factor Intervention Trial						
5,759	442	27.7	3.4	46.5	6.0	59.3	None			Special Intervention group
5,800	491	27.7	3.5	46.4	6.0	58.9	Quadratic	26.4	1.64	Usual Care group

* Weight (kg)/height (m)^2.
† SD, standard deviation; BMI_{min}, body mass index of minimum mortality.

It is widely accepted that controlling for smoking when examining the BMI-mortality relation is necessary to achieve meaningful results in epidemiologic studies (24). Controlling for smoking makes sense on a priori grounds. However, to our knowledge, no one has systematically examined whether control for smoking actually makes a difference when relating BMI to mortality. We examined whether controlling for smoking does, in fact, change results when analyzing the relation between BMI and mortality. To conduct this investigation, we reconstructed the analyses for our 42 cohorts while ignoring smoking status. We found that this omission leaves the qualitative results unchanged (table 3). The estimates of BMI_{min} do change, however (figure 1): For the 30 cohorts fitting the quadratic model, the random effects (25) average BMI_{min} is 24.6 (standard error (SE) 0.33) when smoking is included, and it increases to 25.3 (SE 0.33) when smoking is excluded. This examination does not address the question of whether the quantitative shape of the curve is altered when smoking is included in the model; however, on the basis of these analyses, we conclude that controlling for smoking status changes the results and therefore data should be controlled for smoking when the BMI-mortality relation is being analyzed.

TABLE 3. Proportional hazards models selected using different models and different subsets of 42 analytic cohorts

Relation selected	Model* 1	Model* 2	Model* 3	Model* 4
None	8	11	15	6
Linear, inverse	0	3	0	1
Linear, direct	1	1	2	1
Quadratic	30	27	25	34
Interaction	3	0	0	0

* Model 1: all observations, with smoking included in the model; model 2: smokers only; model 3: nonsmokers only; model 4: all observations, with smoking not included in the model.

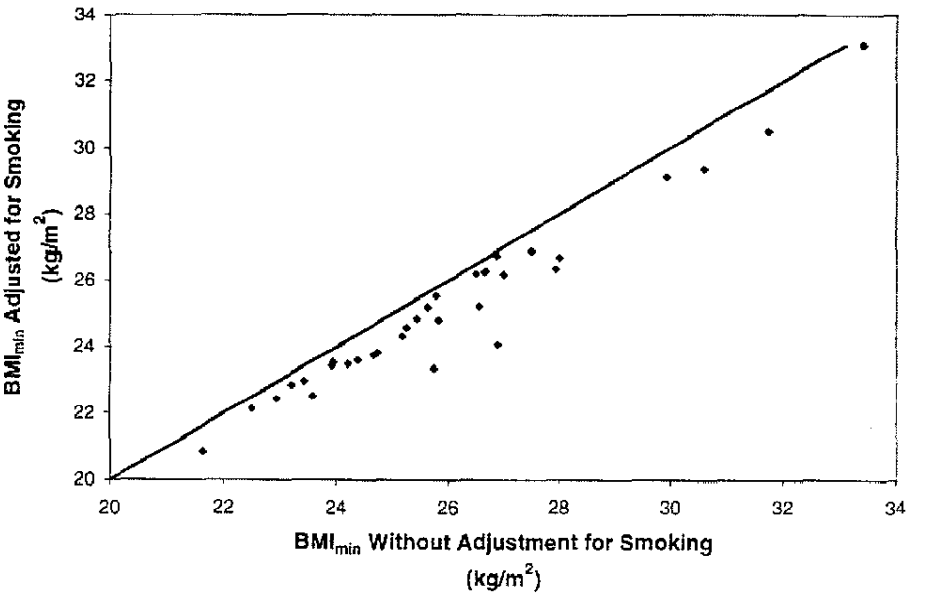

FIGURE 1. Comparison of body mass indices (weight (kg)/height (m)^2) of minimum mortality (BMI_{min}) estimated with and without control for smoking. The average BMI_{min} in 30 cohorts with adjustment for smoking was 24.6 (standard error 0.33); when no adjustment was used, the average BMI_{min} was 25.3 (standard error 0.33). The solid line represents equality.

Table 3 shows the results of reanalyzing the data from each cohort using four different models. The first column of data in the table shows the results of analyses in which the best-supported models were selected for each of the 42 cohorts examined, without deletions. Smoking was included and interaction was tested. The second column presents the analyses of smokers only, and the third the analyses of nonsmokers only. (For the latter two cases, an interaction model was not defined, since it was not possible to include smoking in the model.) The final column presents the results from analysis of the entire cohort, ignoring smoking. We obtained similar results regardless of whether we analyzed nonsmokers or smokers only, and in both cases, few changes were noted.

Figure 2 presents results for the three interaction models examined in these analyses. In each case, a quadratic relation is apparent for both smokers and nonsmokers (see also table 3). In each of the three cases, different quadratic models for the BMI-mortality relation were indicated for smokers and nonsmokers. For National Health Interview Survey white males and Israeli males, the basic shape of the model differed for smokers and nonsmokers but the BMI_{min}'s were similar. In National Health Interview Survey white males, the BMI_{min} was 26.4 (SE 0.23) for smokers and 26.2 (SE 0.49) for nonsmokers. Among Israeli males, the BMI_{min} was 22.4 (SE 1.08) for smokers and 22.4 (SE 0.68) for nonsmokers. For Renfrew and Paisley women, the BMI_{min} was 27.1 (SE 2.20) for smokers and 23.4 (SE 0.76) for nonsmokers.

In table 4, we present the results of a single simulation experiment to illustrate the difficulties involved in excluding observations from cohort analyses. We eliminated set percentages of randomly selected data from each cohort prior to employing the algorithm (i.e., fitting four proportional hazards regression models and determining the one that best fits the data using likelihood ratio statistics). As the percentages of the data exclusions increased from 10 percent to 80 percent, we found an increasing number of changes in comparison with the inclusion of all data in the model. Since these deletions were random, the changes achieved could have no biologic significance. If the elimination of smokers is a unique requirement for analysis, we should obtain results discernibly different from those produced by randomly deleting a proportion of the sample equal to the number of smokers prior to analysis. If such discernible differences do not occur—i.e., if we obtain the same results when randomly deleting a proportion of the sample equal to the number of smokers within the sample—then there is nothing significant about the nature of the data excluded.

We conducted 1,000 repetitions of the simulation experiment for all 42 cohorts. For each repetition, we eliminated a percentage of randomly selected data prior to analysis and then selected the most appropriate model using our standard algorithm. We repeated

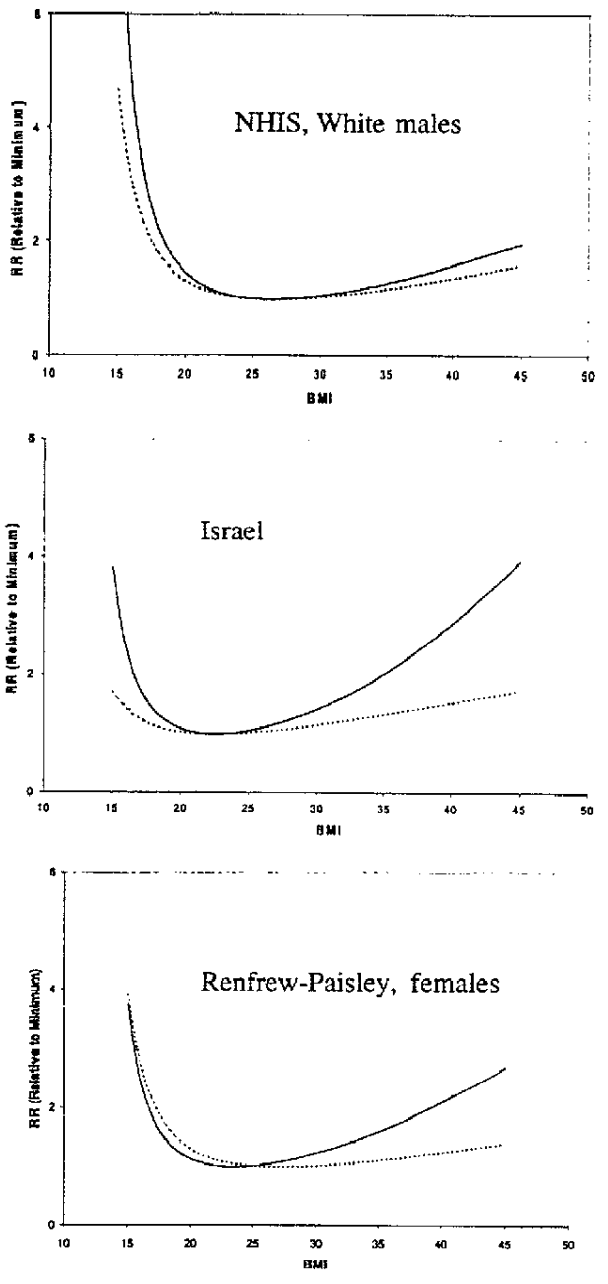

FIGURE 2. Log relative risk (RR) of mortality by body mass index (BMI) (weight (kg)/height (m)^2) in models estimating the interaction between smoking and the BMI-mortality relation in three cohorts. Solid line (—), nonsmokers; dashed line (- - -), smokers. NHIS, National Health Interview Survey.

this process 1,000 times and tabulated the number of cases in which the model selected for the reduced data set differed from that selected when the complete data set was analyzed.

TABLE 4. Results (number of cohorts) of randomly deleting a specified percentage of observations prior to analysis (n = 42 cohorts)

% deleted	Relation selected				
	None	Inverse	Direct	Quadratic	Interaction
0	8	0	1	30	3
10	9	0	1	30	2
20	11	0	1	28	2
30	11	0	1	27	3
40	14	1	1	24	2
50	14	2	1	21	4
60	13	2	1	24	2
70	19	0	2	19	2
80	21	1	3	16	1

Table 5 compares the outcomes from the randomization experiment with those from the analysis of the BMI-mortality relation involving the complete data sets. The top half of the table compares results obtained using the entire cohort with results obtained after excluding smokers from the analysis. The bottom half of the table presents the results of the randomization experiment. For ease of comparison, the numbers in the lower half of the table have been divided by 1,000. For example, among the 8,000 randomized simulations conducted on the eight cohorts that indicated no relation when all of the data were included in the analysis, 6,134 agreed with this original finding, 1,027 resulted in inverse relations, 42 resulted in direct relations, and 797 resulted in quadratic relations.

DISCUSSION

Although health professionals acknowledge the adverse effects of obesity on health, there is still much discussion concerning optimal body weight, defined as the weight an individual should maintain to maximize wellness and life expectancy. Most observational studies show a high mortality rate for both the leanest and the heaviest groups; i.e., there exists a nonmonotonic relation between BMI and mortality. Explanations for the finding of a high mortality rate among the lean vary (26), but some investigators attribute it partly to smoking. Since smokers tend to be leaner than nonsmokers and have a higher mortality rate, it is suggested that the nonmonotonic relation between BMI and mortality might be eliminated if data were controlled for smoking. To our knowledge, this widely held belief has not been tested previously.

We examined the effect of smoking on the BMI-mortality relation using our collection of heterogeneous studies, which contained national samples, cohort studies, and clinical trials. We cannot claim that our data set represents a complete—or even a random—

TABLE 5. Results (number of cohorts) obtained after deleting smokers (top) and results obtained from 1,000 simulations after randomly deleting the same number of observations as there were smokers in the cohort (bottom)

Results with full cohort	Relation selected				Total
	None	Inverse	Direct	Quadratic	
Results using only nonsmokers					
None	7	0	0	1	8
Direct	0	0	1	0	1
Quadratic	8	0	1	21	30
Interaction	0	0	0	3	3
Total	15	0	2	25	42
Results (divided by 1,000) after random deletions					
None	6.134	1.027	0.042	0.797	8
Direct	0.676	0.000	0.295	0.029	1
Quadratic	4.275	2.296	0.121	23.308	30
Interaction	0.010	0.000	0.004	2.986	3
Total	11.095	3.323	0.462	27.120	42

collection of all of the available studies, though all major study designs were included within the sample. Some investigators declined our invitation to participate in the collaboration, and in one case the investigators withdrew their study after the analyses were conducted. However, we have no reason to assume that our results would change with the inclusion of additional studies in our comprehensive data set. We also believe this to be the only analysis to date to have subjected multiple studies to identical analytic procedures.

We conducted several examinations of sensitivity in our results to validate our methodology. First, we adopted an algorithm, which we applied mechanically to the data from each cohort, to select the most appropriate model. We selected as part of this algorithm the p value 0.05 and repeated the analyses using $p = 0.10$. The greater value did not alter the outcome substantially, nor did it change considerably when we added past smokers, in addition to current smokers, to our definition of smoking. Finally, we repeated the analyses using logistic regression rather than the proportional hazards models, with few changes being noted.

Epidemiologic studies have reported the association between BMI and mortality to be positive (27), J-shaped (28), inversely J-shaped (29), U-shaped (30), nonexistent, and even inverse (31). We conducted these analyses to determine whether we could reproduce those conflicting results. We found, however, that when we subjected these studies to a uniform analytic approach, the results were surprisingly consistent. A quadratic relation was found in 30 of the analytic cohorts, and in only three was an interaction between smoking and BMI discerned.

The question is whether the interaction models provide evidence that the shape of the basic relation between BMI and mortality differs for smokers and nonsmokers in these cohorts. There appears to be a widely held belief that if smokers are eliminated from a data set prior to analysis, a direct relation between BMI and mortality will be observed. Because of this view, we have provided an in-depth presentation of the interaction models.

Our findings raise the issue of whether the common practice of eliminating smokers prior to analysis of the BMI-mortality relation is justified. In our view, strong arguments can be made against this practice. The number of smokers within a cohort is often large, and deletions that may constitute more than half of the sample result in loss of power to discern true relations. In addition, with the sample having been diminished to such an extent, confusion results concerning the proper population to which the analytic results apply. Smokers themselves vary from study to study and do not necessarily share a set of common characteristics, which makes these eliminations from analysis problematic.

In our simulation study, we demonstrated that excluding smokers from a BMI-mortality analysis produces findings similar to those resulting from an equal number of random exclusions from the data set. On the basis of these results, we caution investigators against subsetting data without applying formal statistical procedures to test whether the omissions are biologically meaningful. If the kinds of random eliminations we performed above produce outcomes that differ from those resulting from the deletion of the specified data, there is evidence that such omissions are biologically meaningful and hence justified. Otherwise, regardless of widely accepted beliefs that certain subsets (such as smokers) will confound results, the omissions cannot be justified.

The practice of excluding smokers rests on the knowledge that smoking adversely affects mortality risk and

that smokers are known to be leaner than nonsmokers, on average. On the basis of these facts, some researchers would reason that smoking may account for the upturn in mortality at the lower end of the BMI distribution. We sympathize with the initial attractiveness of this reasoning. However, when one breaks data into subsets, the results will often differ among the parts as compared with the whole or even compared with each other. What, then, should be a guide for drawing inferences? Should we simply pick the one we believe or like, particularly if it matches our preconceptions, or should there be some formal criteria for deciding when the subgroup findings are important? We believe that standard criteria should be used to make this decision, and here we introduced the randomization test as a possible standard against which to measure results derived from deletion. If deliberate deletions produce results that are clearly discernible from random deletions, the deliberate deletions may be a valid basis for inference; but when random deletions produce analytic results similar to those from a deletion based on a particular characteristic, we think that the inference based on the deletion is questionable.

In conclusion, our statistical tests showed no evidence of a universal interaction between smoking and BMI that affects the BMI-mortality relation. We found that while controlling for smoking does change the estimate of the BMI_{min}, it does not substantially alter the basic shape of the relation. Finally, we demonstrated that the results obtained by eliminating smokers from the analyses do not differ significantly from results obtained by making random deletions from the analytic data set. The belief that smoking is responsible for the quadratic relation between BMI and mortality or that it explains the excess of mortality among the leanest groups is not supported by empirical observation or quantitative testing.

ACKNOWLEDGMENTS

This work was partially funded by grant DK52329 from the National Institutes of Health.

Investigators: *The Tecumseh Community Health Study*—Dr. Victor Hawthorne; *The Yugoslavia Cardiovascular Disease Study*—Drs. Djordje Kozarevic and Nikola Vojvodic; *The Renfrew and Paisley Survey and The Scottish Collaborative Study*—Dr. Charles R. Gillis, Dr. David Hole, Carole Hart, and Dr. Victor Hawthorne; *the Glostrup Cohort*—Drs. Torben Jorgensen, Troels Thomsen, and Michael Davidson; *The Israeli Ischemic Heart Disease Study*—Dr. Uri Goldbourt and Shlomit Yaari; *Loyola University Medical Center (Coordinating Center)*—Guichan Cao and Drs. Richard Cooper, Ramon Durazo-Arvizu, Zhaohi Li (present affiliation: George Washington University), Youlian Liao, Daniel McGee, T. Elaine Prewitt, and Christopher Sempos (present affiliation: Office of Minority Health, National Institutes of Health).

Data from the following studies were obtained from the National Heart, Lung, and Blood Institute: the Framingham Heart Study, the Hypertension Detection and Follow-up Program, the Lipid Research Clinics Program Prevalence Study, the Multiple Risk Factor Intervention Trial, the Puerto Rico Heart Health Program, and the Honolulu Heart Program. Data from the First and Second National Health and Nutrition Examination Surveys and the National Health Interview Survey were obtained from the National Center for Health Statistics.

REFERENCES

1. NHLBI Expert Panel on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults—the evidence report. Obes Res 1998;6(suppl 2):51S–209S.
2. Massey J, Moore T, Parsons V, et al. Design and estimation for the National Health Interview Survey, 1985–94. (Vital and Health Statistics, series 2, no. 110). Washington, DC: US GPO, 1989. (DHEW publication no. (PHS) 89-1354).
3. National Center for Health Statistics. Plan and operation of the NHANES I Epidemiologic Followup Study, 1987. (Vital and Health Statistics, series 1, no. 27). Washington, DC: US GPO, 1992. (DHEW publication no. (PHS) 92-1303).
4. McDowell A, Engel A, Massey JT, et al. Plan and operation of the Second National Health and Nutrition Examination Survey, 1976–80. (Vital and Health Statistics, series 1, no. 15). Washington, DC: US GPO, 1981. (DHHS publication no. 81-1317).
5. Dawber T, Meadors G, Moore F. Epidemiological approaches to heart disease: The Framingham Study. Am J Public Health 1951;41:279–86.
6. Garcia-Palmieri MR, Costas R Jr, Cruz-Vidal M, et al. Risk factors and prevalence of coronary heart disease in Puerto Rico. Circulation 1970;42:541–9.
7. Epstein FH, Ostrander LD, Johnson BC, et al. Epidemiological studies of cardiovascular disease in a total community—Tecumseh, MI. Ann Intern Med 1965;62:1170–87.
8. Kozarevic D, Pirc B, Dawber TR, et al. Prevalence and incidence of coronary disease in a population study: The Yugoslavia Cardiovascular Disease Study. J Chronic Dis 1971; 24:495–505.
9. Kagan A, Harris BR, Winkelstein W Jr, et al. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: demographic, physical, dietary and biochemical characteristics. J Chronic Dis 1974; 27:345–64.
10. Smith GD, Hart C, Blane D, et al. Lifetime socioeconomic position and mortality: prospective observational study. BMJ 1997;314:547–52.
11. Isles CG, Hole DJ, Gillis CR, et al. Plasma cholesterol, coronary heart disease, and cancer in the Renfrew and Paisley survey. BMJ 1989;298:920–4.
12. Groen JJ, Medalie JM, Neufeld HN, et al. An epidemiologic investigation of hypertension and ischemic heart disease within a defined segment of the adult male population of Israel. Isr J Med Sci 1968;4:177–94.
13. Agner E. Some cardiovascular risk markers are also important in old age. Acta Med Scand Suppl 1985;696:3–50.
14. Hagerup LM. Coronary heart disease risk factors in men and women: from the population study in Glostrup, Denmark. Acta Med Scand Suppl 1974;557:1–116.
15. Hollnagel H. The health structure of 40-year-old men and women in the Glostrup area, Denmark—an epidemiological survey. General design, sampling results and referrals for fur-

ther medical care. Dan Med Bull 1980;27:121–30.
16. WHO MONICA Project Principal Investigators. The World Health Organization MONICA Project: a major international collaboration. J Clin Epidemiol 1988;41:105–14.
17. Lipid Research Clinics Program Epidemiology Committee. Plasma lipid distributions in selected North American populations: The Lipid Research Clinics Program Prevalence Study. Circulation 1979;60:427–39.
18. Hypertension Detection and Follow-up Program Cooperative Group. The Hypertension Detection and Follow-up Program. Prev Med 1976;5:207–15.
19. The Multiple Risk Factor Intervention Trial Research Group. Mortality rates after 10.5 years for participants in the Multiple Risk Factor Intervention Trial: findings related to a priori hypotheses of the trial. JAMA 1990;263:1795–801.
20. Durazo-Arvizu R, McGee D, Li Z, et al. Establishing the nadir of the body mass index-mortality relationship. J Am Stat Assoc 1997;92:1312–19.
21. Cornfield J, Gordon T, Smith W. Quantal response curves for experimentally uncontrolled variables. Bull Int Stat Inst 1961; 38:97–115.
22. Nevill AM, Holder RL. Body mass index: a measure of fatness or leanness? Br J Nutr 1995;73:507–16.
23. Edgington ES. Randomization tests. 2nd ed. New York, NY: Marcel Dekker, Inc, 1987.
24. Manson JE, Willett WC, Stampfer MJ, et al. Body weight and mortality among women. N Engl J Med 1995;333:677–85.
25. DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clin Trials 1986;7:177–88.
26. Manson JE, Stampfer KJ, Hennekens CH, et al. Body weight and longevity: a reassessment. JAMA 1987;257:353–8.
27. Garrison RJ, Feinleib M, Castelli WP, et al. Cigarette smoking as a confounder of the relationship between relative weight and long-term mortality: The Framingham Heart Study. JAMA 1983;249:2199–203.
28. Lew EA, Garfinkel L. Variations in mortality by weight among 750,000 men and women. J Chronic Dis 1979;32:563–76.
29. Menotti A, Descovich GC, Lanti M, et al. Indexes of obesity and all-causes mortality in Italian epidemiological data. Prev Med 1993;22:293–303.
30. Schroll M. A longitudinal epidemiological survey of relative weight at age 25, 50 and 60 in the Glostrup population of men and women born in 1914. Dan Med Bull 1981;28:106–16.
31. Tyroler HA, Knowles MG, Wing SB, et al. Ischemic heart disease risk factors and twenty-year mortality in middle-age Evans County black males. Am Heart J 1984;108:738–46.

APPENDIX

Studies Included in the Analysis

The National Health Interview Survey

The National Health Interview Survey (NHIS) is a continuing nationwide survey of the US civilian noninstitutionalized population conducted through households (2). Data on self-reported weight and height are available for all participants. Specific health topics (supplements) are added each year to the core questionnaire. Information on smoking is available only for the years 1987–1990. Smoking status is classified as never, former, or current. The cancer risk factor supplement provides the smoking data for 1987; the occupational supplement provides the data for 1988; and the health promotion and disease prevention supplement provides the data for 1990. Data on smoking for 1989 are provided in the diabetes supplement; smoking status is coded as 1 (yes) or 0 (no), but the information is available only for half of the adult sample. We recoded smoking status as yes/no for all 4 years by classifying only those persons currently using cigarettes as smokers.

For 1986 onward, linkage information is available on NHIS respondents to allow for matching with other data systems, including the National Death Index. The ability to link NHIS respondents to the National Death Index provides a longitudinal component of the NHIS which allows for ascertainment of vital status. To date, data on multiple causes of death are available for the NHIS survey years 1986–1994, with follow-up through December 31, 1995.

Our analyses included the 121,208 participants for whom information on vital status, age, body mass index, and smoking status was available from the 1987–1990 surveys. These participants contributed four analytic cohorts when stratified by sex and ethnic group (table 2). Information on Hispanic ethnicity was available for NHIS participants. However, the number of deaths among Hispanic participants was too low to be analyzed separately, so for this analysis race rather than ethnicity was used to classify participants. This resulted in most Hispanic participants' being classified as White.

The NHANES I Epidemiologic Follow-up Study

The NHANES I Epidemiologic Follow-up Study data provide follow-up for morbidity and mortality among 14,407 individuals aged 25–74 years initially who received complete medical examinations during the First National Health and Nutrition Examination Survey (NHANES I), conducted from 1971 to 1975 (3). Follow-up surveys were conducted from 1982 to 1984, in 1986 (surveying persons aged ≥55 years at baseline), and again in 1987. For our analyses, we used vital status ascertained through the 1987 follow-up.

The NHANES I Epidemiologic Follow-up Study provided four cohorts and 12,730 participants for our analyses (table 2).

The Second National Health and Nutrition Examination Survey and mortality follow-up

The Second National Health and Nutrition Examination Survey (NHANES II) was conducted from 1976 through 1980 in a nationwide probability sample of approximately 28,000 persons aged 6 months through 74 years from the civilian, noninstitutionalized US population. Baseline data were similar to those of NHANES I (4). A mortality follow-up was conducted for the 9,252 original NHANES II participants aged ≥30 years in December 1992, with over 2,000 deaths being identified.

This study provided four cohorts and 9,064 participants for our analysis (table 2).

The Framingham Heart Study

The Framingham Heart Study was begun in 1948 to investigate factors associated with the development of cardiovascular disease in a representative sample of the adult population of Framingham, Massachusetts (5). A random sample of households was selected, with a response rate of 69 percent